United States
Environmental Protection
Agency

Office of Health and
Ecological Effects
Washington DC 20460

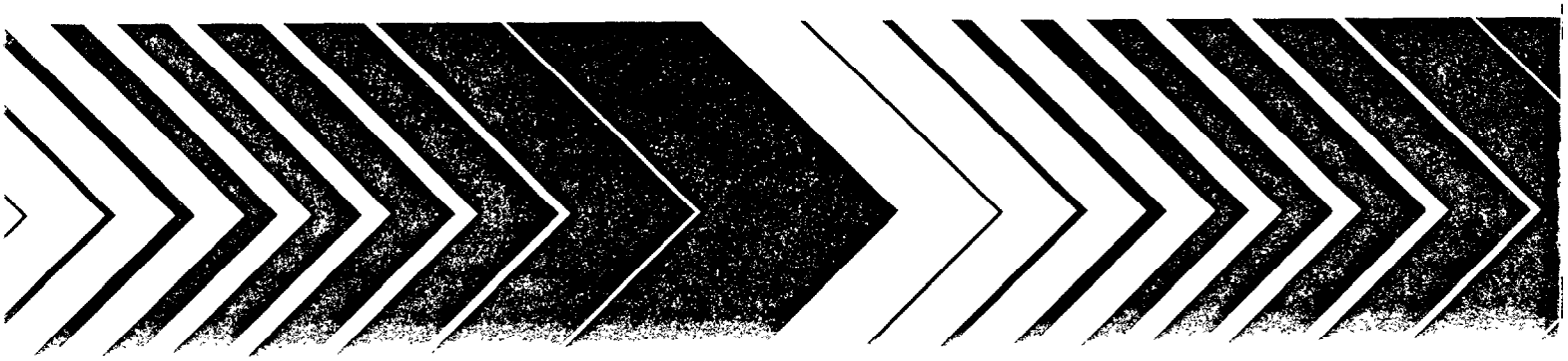
EPA-600/5-79-001a
February 1979

Research and Development

EPA

Methods Development for Assessing Air Pollution Control Benefits

Volume I, Experiments in the Economics of Air Pollution Epidemiology



EPA-600/5-79-001a
February 1979

METHODS DEVELOPMENT FOR ASSESSING AIR POLLUTION CONTROL BENEFITS

Volume I

Experiments in the Economics of Air Pollution Epidemiology

by

Thomas D. Crocker and William Schulze
University of Wyoming
Laramie, Wyoming 82071

Shaul Ben-David
University of New Mexico
Albuquerque, New Mexico 87131

Allen V. Kneese
Resources for the Future
1755 Massachusetts Avenue, N.W.
Washington, D.C. 20036

USEPA Grant #R805059010

Project Officer
Dr. Alan Carlin
Office of Health and Ecological Effects
Office of Research and Development
U.S. Environmental Protection Agency
Washington, D.C. 20460

OFFICE OF HEALTH AND ECOLOGICAL EFFECTS
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

DISCLAIMER

This report has been reviewed by the Office of Health and Ecological Effects, Office of Research and Development, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

PREFACE

The motivation for this volume originated in the authors' mutual and reinforcing convictions that economic analysis and its techniques of empirical application could contribute to the resolution of certain puzzles in studies of the incidence and severity of diseases in human populations, particulary the epidemiology of air pollution. The prior works of Lester Lave, Eugene Seskin, and V. Kerry Smith have provided an excellent base from which to initiate our efforts. These researchers, in addition to Dennis Aigner, Shelby Gerking, Leon Hurwitz, and Roland Phillips have also provided many worthwhile comments and criticisms. None of these individuals are responsible, however, for the results we have obtained.

ABSTRACT

This study employs the analytical and empirical methods of economics to develop hypotheses on disease etiologies and to value labor productivity and consumer losses due to air pollution-induced mortality and morbidity. Since the major focus is on methodological development and experimentation, all the reported empirical results are to be regarded as tentative and on-going rather than definitive and final.

Two experiements have been conducted. First, using aggregate data from sixty U.S. cities, 1970 city-wide mortality rates for major disease categories have been statistically associated with aggregate population characteristics such as physicians per capita, per capita cigarette consumption, dietary habits, air pollution and other factors. Dietary variables, smoking, and physicians per capita are highly significant statistically. However, the estimated contribution the latter variable makes to reducing mortality rates becomes evident only after we recognize that human beings attempt to adjust to disease by seeking out more medical care. The estimated effect of air pollution on mortality rates is about an order of magnitude lower than some other estimates. Nevertheless, rather small but important associations are found between pneumonia and bronchitis and particulates in air and between early infant disease and sulfur dioxide air pollution.

The second experiment, which focused on morbidity, employed data on the generalized health states and the time and budget allocations of a nationwide sample of individual heads of household. For the bulk of the dose-response expressions estimated, air pollution appears to be significantly associated with increased time being spent acutely or chronically ill. Air pollution, in addition, appears to influence labor productivity, where the reduction in productivity is measured by the earnings lost due to reductions in work-time. The reduction in productivity and to air pollution-induced chronic illness seems to be much larger than any reductions due to air pollution-induced acute illness.

CONTENTS

CONTENTS
(continued)

FIGURES

TABLES

TABLES
(continued)

TABLES
(continued)

CHAPTER I

INTRODUCTION TO VOLUME I

Volume I focuses on developing methodology for valuing the benefits to human health associated with air pollution control. Air pollution may affect human health in three ways: (1) by increasing mortality rates, (2) by increasing the incidence and the severity of chronic illness (morbidity), and (3) by increasing the incidence and the severity of acute illness (morbidity).

A number of approaches for determining health effects and valuing them in economic terms are developed within the study. First, if a dose-response relationship is known between mortality rates and air pollution or between days lost from work due to illness (productivity loss) and air pollution, economic losses can be approximated. In the former case, one must know how consumers value increased safety. Thus, if air pollution control reduces risk of death from air pollution related disease, studies of the value consumers place on safety in other situations -- on the job, in transportation, etc. -- can be applied to measuring, the benefits of pollution control programs. Note, however, that valuing safety for small changes in risk is very different from the alternative of valuing human life through lost earnings -- an approach rejected here. Rather, the focus is on examining the value of safety to individuals; that is, how much consumers are willing to pay for safety obtained through pollution control. For morbidity losses, lost time from work and lost productivity during hours of work can be relatively easily valued using observed wage rates.

A second approach for valuing the effects of air pollution on human health is to attempt to observe the effect of air pollution directly on economic factors, thus avoiding the necessity of developing dose-response relationships. If one can develop relationships employing data on wages, wealth, socioeconomic and health status characteristics as well as air pollution concentrations, consumer willingness to pay to avoid illness can be derived. We term this second methodology the willingness to pay approach. It is based on traditional microeconomic theory.

Volume I contains two experiments. First, a data set on sixty U.S. cities is explored to determine if some of the problems of aggregate epidemiology -- epidemiology using aggregate data on groups of individuals as opposed to data on individuals -- can be overcome. The study attempts to estimate a human dose-response function wherein city-wide mortality rates for major disease categories in 1970 are statistically related to population characteristics such as doctors per capita, cigarettes per capita,

information on dietary patterns, race, age and air pollution. The study is unusual in two respects. First, it is the first such aggregate epidemiological study of the effect of air pollution on mortality to include dietary variables, which, along with smoking and medical care, prove to be highly significant. Second, it may be the first study using aggregate data to account for the fact that human beings will attempt to adjust to disease by seeking out more medical care. Thus, cities with high mortality rates are likely to have more doctors per capita. This adjustment process has in the past prevented an estimate of the direct effect of doctors on the prevention of disease. An estimation technique for handling this bias problem is employed, which identifies the contribution medical care makes in reducing mortality rates. The impact of including these new variables in the analysis is substantial.

The second experiment focuses on morbidity rather than mortality. It employs data on the health and the time and budget allocations of a random sampling of the civilian population nationwide. The sample, which was collected by the Survey Research Center of the University of Michigan, consisted of approximately 5,000 heads of households for nine years from 1967 through 1975. Generalized measures of acute illness, stated in terms of annual work-days ill, and of chronic illness, stated in terms of years ill, are available.

The procedures used to estimate dose-response expressions have two somewhat unusual features: (1) care has been taken to employ as explanatory variables only those factors not influenced by the individual's current decisions or health status; and (2) by randomly drawing different samples of individuals, substantial effort was devoted to replicating results.

This volume begins in Chapter II by discussing the role of economic analysis in epidemiology. We then introduce in Chapter III the formidable list of statistical problems faced by epidemiological analysis of air pollution. Finally, Chapters IV and V present the Sixty-City and Michigan Survey Experiments, respectively. Chapter VI presents additional economic results on the valuation of air pollution-induced morbidity.

Chapter II

SOME ISSUES

2.1 Epidemiology and Economics

The motivation for this volume originated in the authors' mutual and reinforcing convictions that economic analysis and its techniques of empirical application could contribute to the resolution of certain puzzles in studies of the incidence and severity of diseases in human populations, particularly the epidemiology of air pollution. The results of our initial efforts to provide empirical support for this perspective are presented in succeeding chapters. Before proceeding to these chapters, however, it is necessary, in order to display the basic rationale for our empirical efforts, to explain our position that economics has some worthwhile things to offer epidemiology.

Many reviews of the epidemiological literature dealing with pollution have remarked upon the relative lack of consistent findings across studies for the effects thought to be caused by any one pollutant. Various reasons are typically advanced for these inconsistencies: inadequate characterization of the pollutants; the use of noncomparable, and sometimes questionable, estimating techniques; failure to account for other environmental influences and self-induced health stresses such as ambient temperature and cigarette smoking; failure to distinguish between pollution levels at work and at home; insufficient attention to differences in genetic endowments, and other factors. The list is sufficiently long and repetitive to be reminiscent of the beat of a somber military cortege. This march has two elements: measurement error and specification error.

The first error element refers to the fact that some variables included in epidemiological studies are inaccurately measured. Sources of error of this sort, however, are hardly unique to epidemiology. They are at least equally common in empirical applications of economic analysis and will therefore be accorded our scrutiny when we discuss our empirical efforts. For the moment, we wish to consider those possible sources of specification error in epidemiological studies that have a basis in the microeconomic theory of the behavior of the individual human being. Our fundamental point is that human beings, the objects of epidemiological attention, make purposive choices with respect to health states and phenomena that influence health states. To the extent that health states are a result of the individual's purposive acts, one must explain these acts in order to comprehend the determinants of the health state. Microeconomics provides a means for grasping the determinants of the individuals's purposive acts.

Acceptance of this perspective adds another dimension (in addition to the social provision of preventive and ameliorative medical inputs) by which social policy can influence the health states of the population, i.e., those factors that influence choices of acts affecting health states can serve as policy instruments.

Specification error occurs in epidemiology (and in economics) when some varibles relevant to the explanation of variations in the health state of interest are improperly introduced or are altogether excluded from the analysis. The biased and incosistent estimates that are the result of excluding nomorthogonal explanatory variables from an expression to be estimated are well-known and intuitively obvious in any case. One can hardly, for example, expect to obtain an accurate estimate of the impact of cigarette smoking on circulatory diseases if the ages of the sample individuals are not controlled. Less obvious, however, are the reasons why common economic variables such as prices often are relevant to epidemiological analyses and why certain variables, both biologic and economic, are sometimes improperly introduced to these analyses.

Some of the most widely known findings in the epidemiology literature concern the respiratory effects (cancer, acute bronchitis, emphysema, the common cold, and pneumonia) of air pollution. View the absence of these respiratory effects as an output that can be reduced by various combinations of clean air and ameliorative medical care, where the latter are considered to be inputs. The literature suggests that there are significant differences in the input-input ratios and in the input-output ratios among various locales, where these locales frequently differ in population size. Suppose it has been observed that:

1. Per capita absence of respiratory diseases is inversely associated with city size.

2. Per capita availability of ameliorative medical care is directly associated with city size.

3. Per capita absence of respiratory disease is directly associated with per capita availability of clean air and ameliorative medical care.

4. Per capita clean air is inversely associated with city size.

5. Respiratory disease absence per unit of clean air and ameliorative medical care is directly associated with city size.

Do the five observations have sufficient informational content to justify a judgment that the dirty air often found in large population concentrations is associated with greater incidence of respiratory diseases and is therefore a plausible cause of these diseases? It would not be surprising if different epidemiological investigators drew a variety of largely contradictory conclusions about the relationships between respiratory diseases, clean air, and ameliorative medical care from these five observations. Contradictions are perhaps inevitable because the ratios expressed in the observations will often be inappropriate means by which to attempt to make judgments about

the relative susceptibilities of human beings to respiratory diseases.

An intuitive notion of the incidence of a disease refers to the frequency of occurrence, given particular levels of instigating factors. Intuition is sometimes misleading. Observation (1) suggests that small cities have less incidence because they have less respiratory disease. Observation (5) leads to the opposite conclusion since large cities have fewer respiratory diseases relative to their clean air. But observation (4) makes small cities look favorable because of their greater provision of clean air. Or do large cities subject their populations to greater incidence of respiratory effects by having fewer units of ameliorative medical care available? Observation (3) again favors small cities because of the greater per capita availability of ameliorative medical care.

One might suspect from (5), (4), and (2) that larger cities have more ameliorative medical care relative to clean air than do smaller cities. The former have dirtier air and thus try to compensate by providing additional ameliorative medical care. It is thus not surprising that the ratio of of absence of disease per unit of available medical care favors the larger cities. An alternative interpretation of (3) is that disease frequency increases with city size not only because of dirtier air but also because the price to the consumer of medical care is greater than in smaller cities. Greater prices of these services for the consumer can imply greater returns for the profession that provides these services. Greater returns attract these professionals, resulting in greater availability of their services. However, these same higher prices also mean that sufferers from a respiratory disease of given severity will seek out less ameliorative medical care. Are then these prices, the dirty air, or the consumption of medical care the causes of the incidence of the respiratory disease? Recognition that they are intertwined is a significant but insufficient step. The nature of the intertwining remains to be explained.

2.2 When Microeconomics Doesn't Matter

Microeconomic analysis specifies the conditions under which decisionmakers (human beings) are expected to have identical ratios of inputs and outputs. Basically, these identical ratios would occur if: (1) all decisionmakers had identical biological endowments and transformed inputs into health states in precisely the same fashion; (2) all decisionmakers faced the same prices in (implicit and explicit) input and health state markets; (3) all decisionmakers had the same real income; and (4) all decisionmakers had identical preference orderings. If all these conditions were fulfilled with respect to a particular pollutant, only one point could be observed on the epidemiologist's dose-response curve: there would be no variation whatsoever in the observable behavior of individuals.

We nevertheless observe decisionmakers in the real world with similar states-of-health who have different biological endowments and varying ways of transforming inputs into these health states. One can, of course, pass muster in explaining the real world by assuming that decisionmakers (?) behave randomly or that all health states, whether present or future, are determined by physical or biological factors beyond the decisionmaker's present control.

This is no different than assuming that the decisionmaker is abysmally ignorant of cause-and-effect with respect to health states or that he just does not care about his health state. If any of the conditions in this paragraph are in fact true, then current epidemiological precedures, which tend to give short shift to economic variables and which implicitly treat the individual as being completely unable to exercise influence over events that affect his choices, are entirely satisfactory. This abrupt statement requires clarification.

Panels I through VI of Figure 2.1 represent two unit isoquants (loci of points showing all combination of two inputs that will yield equal health states) for inputs X_M and X_A (e.g., medical care and clean air), with the current positions of decisionmakers R (a rural person) and C (a city person) indicated. Each isoquant represents the same state-of-health as the other isoquant. Note that the effectiveness of each input in providing the unit health state for each individual is assumed to decline progressively as more of one input is substituted for the other. Thus additional medical care becomes progressively less effective as the air becomes dirtier. Similarly, cleaner air becomes an increasingly poor substitute for medical care as less and less medical care becomes available.

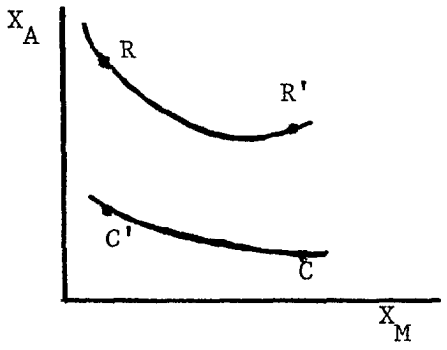
All panels are drawn so that on the basis of his state-of-health per unit of clean air, decisionmaker C is in better shape than decisionmaker R. Conversely, decisionmaker R does better than C in terms of his health state per unit of medical care. In each panel, therefore, C uses relatively less clean air and R uses relatively less medical care to attain the unit health state. This situation is consistent with the previous five observations on the associations between city size, clean air, and ameliorative medical care.

Panels I and II refer to the case where the question of whether economic variables should be included in dose-response function analysis, and, if included, how to include them, need never arise. The clean air and medical care each individual requires to attain the unit health state are determined by physical and biological (technical) considerations alone. Purely economic considerations play no part. Nevertheless, the two panels do provide insights about cautions to exercise when attempting to establish dose-response functions by studying several individuals at one instant in calendar time. In Panel I, in the absence of knowledge about the isoquants of R and C, any attempt to establish the population dose-response function by averaging over the current positions of R and C is doomed to be a misrepresentation. The unit isoquants of Panel I belong to dose-response functions that differ not only by a constant term but which also embody entirely different responses of health states to particular cominations of medical care and clean air. The "average" or population dose-response function or isoquant established by pooling a single medical care-clean air combination from each isoquant will differ according to where each individual happens to be on his isoquant when he is observed. For example, the average of R and C' differs substantially from the average of R' and C. If and only if several medical care-clean air combinations for each individual were observed could a representative dose-response function be obtained. This would generally require that several observations over time be made of each individual.

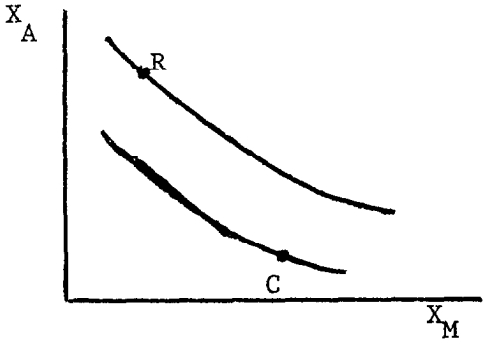
Figure 2.1

Alternative Measures of Disease Incidence

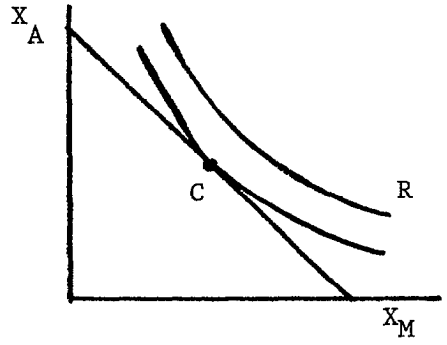
I



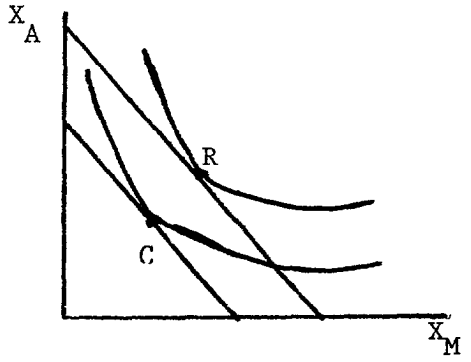
II



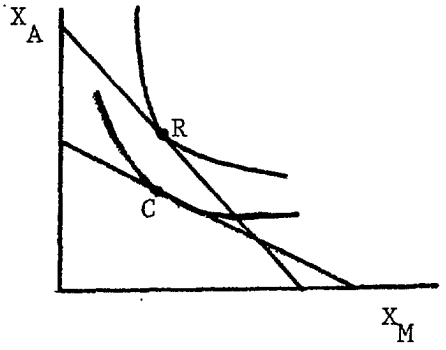
III



IV



V



In Panel II, several observations of each individual over time are not required because the isoquants belong to dose-response functions differing only by a constant term. This term could represent differences in biological endowments, childhood environment, previous lifestyles, and other factors with which epidemiologists traditionally deal. These same factors, however, could also explain the nonconstant difference between the isoquants of Panel I. Clearly, the current situation favors individual C in Panels I and II since he is able to attain the unit health state with smaller quantities of both medical care and clean air.

Panel III introduces the economic information of relative prices and the income that each individual has already decided to devote to health maintenance. Assume, for the moment, that each individual has decided to devote the same income and faces exactly the same prices for medical care and clean air. The result is that individual R is unable to attain or maintain the unit health state, although individual C, given his income and the relative prices, is fully able to do so. Individual R, due to his economic circumstances and his dose-response function, must settle for something less than the unit health state. Both biological and economic factors inhibit him from reaching the unit health state. Insofar as health states do not affect incomes and relative prices, this panel would appear to justify the common epidemiological practice of introducing incomes into a dose-response expression that is to be estimated. Panel IV, which has the incomes of the two individuals differing but presumes they continue to face identical relative prices, also seems to justify this practice. The justification is a mirage.

If the objective of epidemiological investigation is to ascertain the extent to which various physical and biological factors contribute to differences in the R and C-isoquants, then the introduction of income into a dose-response expression must reduce the estimated impact of inputs such as the medical care and clean air of Panel IV. The introduction of income is redundant. Income, along with relative prices and the form of the isoquants, determines the quantities of medical care and clean air each individual consumes. As the panels indicate, for given relative prices, the greater the individuals's income, the more health care and clean air he will consume, assuming he has not yet reached the unit health state. The quantities of medical care and clean air that enter the dose-response function estimate are thus partially determined by each individual's income. Thus the latter is a measure of the former and must capture part of the influence that would and should otherwise be attributed to clean air and medical care. Bluntly, epidemiological studies that include income reduce the odds that clean air will be seen as contributing to good health. The degree to which this reduction in odds is worthy of concern is dependent upon the extent to which income determines the consumption of clean air. The little evidence that is available indicates that at least within individual cities the association between income levels and cleaner air tends to be quite high.

Panel V depicts a situation where individuals R and C have nothing in common: they have different unit health isoquants, devote different income levels to health maintenance, and face different relative prices for medical care and clean air. Both individuals consume similar quantities of medical care but radically different quantities of clean air. Again, however, the

epidemiologist interested solely in dose-response functions can safely neglect giving any attention to incomes and relative prices, for these serve only to determine the quantities of medical care and clean air consumed that directly determine health states. Nevertheless, this conclusion does not justify appealing to observations similar to those mentioned in the previous section as grounds for judging that clean air improves health states.

There are several alternative explanations for the ratios expressed in these observations. Different individuals may have different dose-response functions. Sometimes these differences may be captured by a constant term; at other times, the slopes of the functions may be dissimilar, invalidating attempts to ascertain population dose-response functions solely by observing each sample individual only once. Moreover, variations in individual incomes and in the relative prices of health inputs may be the cause of the observed ratios. This implies that the policymaker can influence the quantities of these health inputs consumed by doing nothing more than manipulating a limited set of purely economic variables. Under the conditions specified in this section, however, these variables have no bearing on estimating, via standard epidemiological procedures, the responses of the human organism to variations in the quantity of clean air.

2.3 When Microeconomics Does Matter

The preceding section employed stated, but not very visible, assumptions to arrive at the conclusion that epidemiological studies err when they devote attention to economic variables in attempting to establish dose-response functions. In particular, it was assumed that the individual had already decided the resources he would dedicate to health maintenance and that this decision did not influence any other decisions he might make. If either or both of these assumptions are inaccurate descriptions of reality, then microeconomics does matter in the determination of dose-response functions. The assumptions had the effect of removing the purposive nature of the human being from consideration: all the individual's choices were presumed to have already been made.

In implicit form, a good approximation of the expressions that epidemiologists frequently use to estimate the response of a particular mortality or morbidity effect to a particular environmental exposure is:

$$\pi_i = f_i(X,Y,Z,E,\varepsilon), \tag{2.1}$$

where π_i is the probability of the ith individual dying or becoming ill from the exposure; X is a vector of available ameliorative medical care inputs; Y is a vector indicating the individuals's socioeconomic class, medical history, ethnic group, etc.; Z is a vector of the individual's activities representing lifestyle habits such as diet and exercise regimens; E is a vector of environmental exposures that, *a priori*, are thought to be physical or biological instigators of the health effect; and ε is a stochastic error.

The form of $f_i(\cdot)$ is typically unknown and must therefore be approximated, perhaps by a linear expression. The coefficient attached to the exposure of interest would, given an acceptable level of statistical significance, then

be interpreted as the increase in the health effect incidence caused by a one-unit change in the exposure. Would it then be reasonable to infer a dose-response association from the coefficient of the exposure variable?

The aforementioned inference would be correct if and only if it is possible to alter the environmental exposure without altering the value of any other explanatory variables in the expression. It is easy to show that this cannot be done when the structure is presumed to consist of no more than one relationship. The reason is that (2.1) contains at least two variables, the current and future levels of which are subject to at least some control by the individual; that is, during the period in which it is thought the health effect can occur, the individual can influence by his voluntary choices the magnitude of explanatory variables supposed to determine the health effect. For example, the probability of the individual suffering the health effect, π, is dependent upon the extent to which he chooses to use the available medical care and the mix and magnitude of activities he chooses to undertake. In order to explain the health effect outcome, one must also explain the structure underlying these choices. The following simple example shows one way in which π and Y, interpreted as income, might be jointly determined.

If both the π and Y functions can be linearly approximated, they can be written as:

$$\pi_i = \alpha_1 + \alpha_2 E + \alpha_3 X + \alpha_4 Y + \alpha_5 Z + \varepsilon_1 \tag{2.2}$$

$$Y = \beta_1 + \beta_2 \pi_i + \beta_3 A + \beta_4 S + \beta_5 L + \varepsilon_2 \tag{2.3}$$

Expression (2.2) states that the question of whether or not the individual is suffering from chronic bronchitis is related respectively to the non-cigarette bronchitis-causing agents (e.g., air pollution) to which he is exposed, the ameliorative medical care he consumes, his income, and the number of cigarettes he smokes. In turn, (2.3) states that the individual's income is determined respectively by whether or not he has bronchitis, his absenteeism rate, his schooling, and the length of time he has been on the job.

Solving (2.2) and (2.3) for π_i alone, we have:

$$\pi_i = \frac{\alpha_1 + \alpha_4 \beta_1}{1 - \alpha_4 \beta_2} + \frac{\alpha_2}{1 - \alpha_4 \beta_2} E + \frac{\alpha_3}{1 - \alpha_4 \beta_2} X + \frac{\alpha_4 \beta_3}{1 - \alpha_4 \beta_2} A + \ldots$$

$$+ \frac{\alpha_5}{1 - \alpha_4 \beta_2} Z + \frac{\alpha_4 \varepsilon_2 + \varepsilon_1}{1 - \alpha_4 \beta_2} \tag{2.4}$$

Consider the coefficient attached to E in (2.4). If E is air pollution, (2.4) shows that an estimate of (2.2) will not yield the response of bronchitis incidence to dosages of air pollution, even though, in the language of epidemiologists, the dose-response is "adjusted" for medical care, life-style, and socioeconomic class. Instead, the coefficient for E in (2.2) will be a mix of effects due to air pollution, income, and the effect of

bronchitis on income. The product of the coefficients for the latter two effects would have to approach zero in order for the response of bronchitis to air pollution alone to be obtained. For this to occur, chronic bronchitis could have no effect on the individual's income and this income could have no effect on his chronic bronchitis. Both assertions, particularly the first, are quite implausible. In fact, in the absence of further information, the sign that would be obtained for the coefficient of E in (2.2) is ambiguous since $\alpha_2 \gtrless 0$, $\alpha_4 \lessgtr 0$, and $\beta_2 \gtrless 0$. It is entirely conceivable, if one were to estimate (2.2) alone, that one would find air pollution reducing chronic bronchitis! In any case, because the product of α_4 and β_2 is negative in sign, the effect of air pollution on health will be underestimated. One could readily obtain a similar result for Z, cigarette smoking.

It might be reasoned that the difficulty with the preceding example could be removed if income were excised as an explanatory variable from (2.2). The expression would not then have any pecuniary variables in it and would therefore seem amenable to the customary epidemiological ministrations. These customary ministrations would, however, continue to be incorrect, for the individual is able to influence the quantity of cigarettes, Z, that he smokes during the current period. If air pollution exposures change, the individual is likely to change the quantity of cigarettes that he smokes. Thus, even after excising the income variable from (2.2), possibilities for biasing the air pollution coefficient remain. To see this, write:

$$\P_i = \alpha_1 + \alpha_2 E + \alpha_3 X + \alpha_4 Z + \varepsilon_1. \tag{2.5}$$

$$Z = \beta_1 + \beta_2 \P_i + \beta_3 P_Z + \beta_4 P_k + \beta_5 Y + \varepsilon_2. \tag{2.6}$$

The variables in expression (2.5) are defined as in (2.2). Expression (2.6) states that the quantity of cigarettes the individual currently smokes is a linear function respectively of whether or not he has chronic bronchitis, the price of cigarettes, the prices of goods that are complements and/or substitutes for cigarettes, and his income.

Upon solving (2.5) and (2.6) for $\P_i$, the coefficient attached to air pollution, E, proves to be $\alpha_2 \left(1 - \alpha_4 \beta_2\right)$, which represents a mix of effects due to air pollution, cigarette smoking, and the effect of bronchitis on cigarette smoking. Again, the product of the coefficients for the latter two effects would have to approach zero for the response of bronchitis to air pollution alone to be obtained. In addition, the sign of the E-coefficient would again be ambiguous since $\beta_2 \gtrless 0$. If $\beta_2 > 0$, the effect of air pollution would be overestimated, and if $\beta_2 < 0$, the effect would be underestimated.

To attempt to account for the additional factors thought to influence a morbidity or mortality rate by simply stringing out variables in a single expression must clearly often be incorrect. During the period in which the health effect is supposed to occur, humans acting in their individual capacities can choose to influence the magnitudes assumed by certain of these

variables. Each variable susceptible to this influence must be explained by an expression of its own. Economic analysis is necessary to impart an interpretable form to these expressions. Physical and biological constructs will therefore often be insufficient tools with which to provide epidemiological explanations of disease incidences.

The previous two examples are about problems of joint determination which involve economic variables. Nevertheless, the problem of joint determination does not require the presence of economic variables. For example, epidemiological studies frequently group disease incidences by individual city and employ measures of central tendency of incidence and other variables as single units of observation. Thus one might try to explain the frequency of deaths from cancer in a sample of U.S. cities by relating it to the dietary habits, air pollution exposures, and median age of the population in each city. Among the dietary variables, one might include saturated fats and cholesterol, dietary components frequently said to be positively related to cardiovascular disease. Inclusion of these two variables in an expression intended to estimate the factors that contribute to cancer incidence would probably result in negative signs being attached to their coefficients, implying that saturated fats and cholesterol prevent cancer. This may, in fact, be true, but only indirectly. Specifically, median age in each city will tend to vary inversely with the incidence of cardiovascular mortality; in other words, earlier death reduces median age. Thus, since cancer incidence is positively influenced by median age, one might expect cancer to exhibit negative associations with saturated fats and cholesterol even if they have no direct causal relationship with cancer incidence. The apparent effects of these two dietary variables upon cancer incidence would actually represent a confounding of: (1) the effect of the two variables upon cardiovascular disease; (2) the relation between cardiovascular disease and median age; and (3) finally, but only via (1) and (2), the effect of the two variables upon cancer incidence. In short, at least one other expression explaining median age is required.

2.4 The Costs of Pollution-Induced Disease

The preceding sections have discussed the circumstances under which microeconomics and its methods of empirical application can contribute to the epidemiology of pollution. It was observed that in trying to establish dose-response functions for particular pollutants, it is necessary to be extremely sensitive to the presence of jointly determined variables. Failure to account properly for these variables in the structure to be estimated can result in badly distorted depictions of the effect of a health input such as pollution upon the output, the state-of-health or the incidence of a particular disease. One could, of course, consider all variables to be endogenously determined in some ultimate sense. The key to stopping short of including the entire universe in the structure to be estimated is the formation of intelligent judgments about those variables important to the question of interest over which the individual or system (e.g., urban areas) can immediately exercise no more than trivial control. The number of expressions must equal the number of variables it is posited that the individual or system can control if a determinant solution is to emerge. Most importantly for our purposes, since many of the jointly determined variables in a dose-response structure will be economic requiring

the application of microeconomic analysis in order to specify how they are to be introduced to the structure, the actual design of epidemiological studies must often include microeconomic considerations.

The potential application of microeconomic analysis to epidemiological concerns extends beyond the estimation of dose-response functions. The analysis can be used to establish pecuniary values for pollution-induced health effects. These values, which are consistent with the axiomatic structure of benefit-cost analysis, can contribute to evaluations of the economic efficacy of existing and proposed pollution control programs. Attempts to establish these values can adopt two polar views of the individual's degree of comprehension of the relation between pollution and his state-of-health.

The first of these views presumes that the individual fails to comprehend any connection between pollution and his health state, even though pollution does influence this state. To obtain the total loss due to a pollution-induced health effect, this view justifies the estimation of a dose-response function and the multiplication of the loss in health attributed to pollution by a pecuniary value for the health loss. The information and criteria used to set the pecuniary value, and thus the total pecuniary loss, come from outside the system being studied. The basic presumption is that the individual is unaware of the health effects of pollution and therefore does not make any voluntary adjustments in response to its presence.

In addition to being a relatively easy and therefore desirable way to establish pecuniary values for health losses, this first view has the further advantage of reducing the force of the joint determination problem. It thus removes problems similar to the cigarette example of the previous section, where, in response to the presence of increased air pollution, the individual chose to reduce his cigarette consumption. However, the view would affect neither the income nor the dietary examples, for the ill-health caused by pollution can affect the individual's earnings capacity and his dietary habits. These earnings and habits would therefore change as pollution changes, even though the individual is utterly unaware of the cause and, consequently, fails to make any behavioral adjustments in response to pollution.

The polar opposite of the above view is that the individual is fully cognizant of the health effects of pollution and continually adjusts his voluntary behavior accordingly; that is, given the opportunities he has available and the relative prices he faces, he alters his behavior so as to minimize the value of the pollution-induced health losses he suffers. These voluntary adjustments will involve shifts in his time and budget allocations such as reductions in the time and intensity of outdoor activities, pursuit of a less toxic diet, and more visits to the family physician. A view of the individual that presumes he is unaware of the health effects of pollution does not account for these adjustments. In effect, it assumes that, whatever the variations in pollution, the individual's time and budget allocations have always accorded with the allocations occurring at the time of observation. Since, according to the second view of the individual's response to pollution variations, these observed

allocations are the result of attempts to mitigate the health effects of pollution, the first view of the individual results in underestimates of pollution health effects. Furthermore, if individuals do reallocate their time and their budgets in response to pollution variations, then measures can be obtained of the income the individual would have to receive or would be willing to pay to leave himself as well off as he was before a change in pollution. These measures correspond to the ideal measures of economic loss established in the microeconomic theory of consumer behavior.

Chapter III

SOURCES OF ERROR

3.1 Problems in Statistical Analysis

The previous chapter introduced the problem of joint determination of many variables - especially those which involve choice by individuals - in epidemiological relationships. This problem, if not explicitly accounted for, can introduce simultaneous equation bias. Estimated effects will not approximate actual (population) values. In other words, even for large samples (those approaching infinity) estimated coefficients are no longer consistent; they do not approach their true population values. A number of techniques are available for providing consistent estimates in simultaneous equations. One of these is described in 4.3 below and the technique is applied both in the Sixty-City experiment, Section 4.5, and in the Michigan Survey experiment, Section 5.6. This chapter thus addresses a number of remaining statistical obstacles to obtaining unbiased estimates and significances of the effects of air quality on human health.

3.2 Heteroskedasticity

Any empirical exercise involves error. To act otherwise is to fool one's self, if not the reader. The error can be due to an inability to capture all the *a priori* factors that influence the phenomenon of interest, it can be caused by failures in measuring the magnitudes of the variables one has *a priori* grounds for introducing, or it may be a consequence of a misunderstanding of the structure of the phenomenon. In addition to altering the estimated values of coefficients and/or confidence intervals, errors are registered in the constant terms and the residuals of estimated expressions. The so-called statistical "classical linear model," which is employed to establish all the relations of this volume, presumes that the mean of the error variance (a measure of the dispersion of the observations of the magnitudes of a variable around its average magnitude) is equal to zero. This implies that the errors are constant for observations on all basic units of analysis.

In our mortality study, if the unexplained portion of the incidence of cancer-induced death tends to increase with the size of city, then the error will not be constant from one observation to another. Similarly, in our morbidity study, if the unexplained portion of the duration of chronic illness increases with the value of some variable, then we have again violated a basic premise of the classical linear model. Thus, for example, one might reasonably expect that in locations where air pollution is low

and that the variation around this average level would not be very great. Low concentrations of air pollution are unlikely to generate severe chronic illnesses of long duration. However, where air pollution concentrations are high, both the average level of air pollution-induced chronic illness and the variations around this average are likely to be substantial. In low pollution locations, even those with a biological propensity to be harmed from pollution do not suffer any ill effects. However, those with this propensity might be struck down if they are moved to a high pollution location, whereas those who have great resistance will suffer little, if at all. The variation in the duration of chronic illness is therefore much higher where pollution is suffocating because the magnitude of the greatest suffering has greatly expanded, while the magnitude of the least suffering continues to be zero.

Nonconstantcy of the variances of the errors (residuals) in an estimated expression is termed "heteroskedasticity," a term the linguistic roots of which we don't know. Because it means that variation in the errors of an expression varies systematically over observations, it implies that the confidence intervals for estimated coefficients will also vary systematically. The result is that the same basis will not be used to calculate the confidence intervals among observations. Thus, although the estimated coefficients are not affected, the standard errors of these coefficients will be biased. As a consequence, the customary tests of significance have no meaning. Nevertheless, if one knows the direction of the bias, one can sometimes ascertain whether these customary tests of significance accord excessive or too little precision to the estimated coefficients. For example, Kmenta (1971, p. 256) provides a formula that under limited circumstances, permits the calculation of this magnitude and the sign of this bias in standard errors. He also outlines ways in which the raw data can be corrected to negate heteroskedasticity.

3.3 Multicollinearity

Multicollinearity occurs when two or more explanatory variables are so highly correlated among themselves that it becomes difficult to separate or determine the independent effect of each variable. In the extreme case where two variables are perfectly collinear, they are effectively identical. However, if a priori information exists on the effect of the collinear variables, then that information can be used. For example, if in attempting to explain stomach cancer mortality rates using cross-sectional data, two explanatory variables, sulfur oxides in air and per capita consumption of Polish sausage, are perfectly collinear, one might employ data from animal experiments or epidemiological studies on select human populations (e.g., Polish populations and industrial workers exposed to SO_2 in high concentrations) to determine the relative incidence of stomach cancer from each factor. By including only one of the variables in the regression, the total effect of both explanatory variables will be captured by the estimated coefficient on that one variable. Thus, if consumption of Polish sausage and sulfur oxide exposures are perfectly collinear and only consumption of Polish sausage is included in the estimated equation, the estimated coefficient on consumption of Polish sausage will capture the effect of both variables. How that effect is to be allocated between the two variables depends on the availability of external information. For example, if animal

experiments do not show a link between sulfur oxide exposures and stomach cancer, but do show a link between consumption of cured meats (including Polish sausage) and cancer, one might allocate the entire coefficient to consumption of Polish sausage. Of course, if this were the case, and the investigator did not know that consumption of Polish sausage and sulfur oxide exposures was perfectly collinear and no dietary data was available for inclusion, then a false link between sulfur oxides and stomach cancer might be shown using the cross-sectional data alone.

The same arguments apply to cases of near perfect multicollinearity wherein explanatory variables are highly, as opposed to perfectly, correlated. This is, of course, the most likely case. However, the outcome of including two or more collinear explanatory variables is an increase in the standard error of the estimated coefficients for the collinear variables. The standard error is, of course, a measure of the accuracy with which a coefficient is estimated -- large standard errors imply that the actual coefficient could be much larger or smaller than the estimated coefficient. Thus, when collinear variables are included, the inability to separate influences is reflected in the measure of uncertainty over the magnitude of the estimated coefficients on those variables.

The approach taken here to deal with multicollinearity -- and the 60-city experiment described below has a severe problem among the dietary variables -- is to a priori exclude variables which are highly collinear with respect to a representative included variable. An alternative approach to multicollinearity is the use of a technique known as ridge regression [see Schwing, et. al. (1974)] which, however, makes interpretation of the resultant estimated coefficients unclear.

While multicollinearity within an available data set makes estimation and interpretation more difficult, at least the problem can sometimes be recognized and false conclusions thereby avoided. However, where unknown collinearity occurs, for example when an included explanatory variable is highly collinear with a variable which is not available to the investigator, the false conclusion can be reached that the included variable is solely responsible for the estimated effect. The investigator may not recognize that the estimated effect includes the effect of one or several other excluded but collinear variables. We discuss this possibility below.

3.4 Causality and Hypothesis Testing

Aside from the problem of multicollinearity, the traditional problems of causality underlying epidemiological studies still apply. For example, if heart attacks are actually related to cigarette consumption, but smoking is correlated with coffee consumption for behavioral reasons, a spurious positive correlation might be shown between heart attacks and coffee consumption, especially if cigarette consumption is excluded from an estimated statistical relationship. In other words, correlation does not prove causation, and statistical hypothesis testing can never confirm, but only reject, a maintained hypothesis. Turning to another example, if most nitrite (used to cure meats) ingestion is through consumption of pork products (70 % of pork is cured), one might suspect, given the hypothesis

of in vivo nitrosamine (a carcinogen) formation from nitrite, that cancer mortality and pork consumption would be correlated. If such a correlation can be shown (as it has been; see Kneese and Schulze (1977) and NAS 1978) then the only valid conclusion is that we do not reject the hypothesis that pork consumption (and perhaps, in turn, nitrite ingestion) is related to human cancer. If, alternatively, one accepts the maintained hypothesis on a priori grounds, and no bias exists in the estimation procedure, regression analysis can give a best linear estimate of the actual relationship in the sample population between, for example; cancer mortality and a dietary factor such as nitrite ingestion. However, regression analysis cannot prove causality; causality must be assumed in this procedure. This is why it is so important to have hypotheses concerning causality before a regression equation is specified.

A set of hypotheses concerning human health, including the effect of air pollution, forms a model of human health. The concept of a complete model of human health as the basis for hypothesis testing is an important one for several reasons. First, a modeling framework immediately suggests that behavioral elements such as voluntary medical care may be important and as pointed out above, a simultaneous equation structure may be necessary to test hypotheses properly. Second, the modeling framework focuses attention on a complete specification of the determinants of human health. A "better" model will exclude fewer relevant variables and be both a more accurate predictor of human health and more accurately identify the effect of each explanatory variable. The modeling approach then helps avoid the problem of unknown collinearity by focusing on a specification which provides imformation about the effects of all relevant variables.

An alternative viewpoint has been expressed by Lave and Seskin (1977). Their argument rests on the assumption that excluded variables (medical care, diet, and smoking are excluded from their study of air pollution and human health) will not bias estimated effects of included variables if the excluded variables are orthogonal (perfectly non-collinear) with respect to the included variables. Thus, if one assumes orthogonality with respect to excluded variables, following Lave and Seskin (1977), one can justify estimation of incompletely specified equations. We take a different approach principally because we reject orthogonality as a reasonable assumption. If, as ecologists are fond of saying, "everything depends on everything else," then simultaneity and collinearity are likely to be pervasive in the "real world." In fact, we argue below based on our own epidemiological and economic data that this is just the case.

Finally, to test specific hypotheses, we will use the standard significance test; we will test the hypothesis that each explanatory variable has no effect (has a coefficient of zero) by using the appropriate t-statistic which, in this case, is approximately equal to the estimated coefficient divided by its own standard error. For example, for large samples, if for a specific coefficient $t \geq 2.0$ (if the coefficient is greater than or equal to twice its own standard error), then, where the hypothesis tested includes an assumed sign for the coefficient, a 97.5% level of significance is achieved. This implies that, in random sampling of a population, one would draw a sample which accidentally confirmed the

hypothesis (effect non-zero) only 2.5% of the time.

It is important to note, that as the significance level is implicitly lowered from t = 2.0 toward t = 1.0, even in large samples, spurious relationships begin not to be rejected. Practical experience and econometric tradition suggest that a 95% to 97.5% significance level is appropriate. The desired confidence level should be chosen a priori to avoid the temptation to "prove" desired relationships by ex post lowering of the level of significance for rejecting or failing to reject hypotheses. Similarly, statements that an explanatory variable is "nearly significant" should be interpreted with great caution. Where costly environmental programs are to be justified by epidemiological analysis, rigorous tests of significance should be employed.

3.5 Aggregation

In one or another of his many books, Herbert Simon has used the term "bounded rationality" with reference to limited human abilities to arrange, comprehend, and manipulate large volumes of information. More succinctly, Simon is referring to the need to simplify in order to understand. Even the pure theorist, in both his analysis and exposition, must partition the universe into two parts: that with which he will and won't deal. Moreover, he must employ a limited and often quite small number of concepts to deal with the part he has chosen. He who would measure as well as theorize must simplify beyond this, for he must be economic with his data manipulations. Both isomorphism with his theoretical variables and his less than fully robust empirical tools require this. Simplification is synonymous with throwing away information, but that which is thrown away is often beyond our powers of use. As Stigler (1967) has remarked, ". . . information costs are the costs of transportation from ignorance to omniscience and seldom can a trader afford to take the entire trip."

In the material to follow, we have played the role of the aforementioned trader in two ways. First, in the mortality study, we have employed grouped data for estimation; that is, we have employed a single measure of central tendency (usually the arithmetic mean) of the distribution of some attribute across a group of people or locations (a city) as the sole representation of the group's diversity. We have melted entire cities into one pot. Here we wish to discuss the issues this poses for estimation.

A second aggregation thing we have done is to embrace the notorious representative individual when discussing the pecuniary benefits or costs of a given health effect. Too fond an embrace of this representative can lead to gross errors if his responses are incautiously applied to flesh and blood individuals. We wish to explain why. Initially, however, we will discuss the estimation issue.

In the mortality study, the unit of analysis is a city or some larger jurisdictional unit and the values attached to a particular variable represent the per capita magnitudes of the variable in the cities. To form these per capita magnitudes, someone had to collect observations on the values of the variables for the distinct individuals in each city. By using the per capita rather than variation of the individual observations within each city

and thereby reducing the efficiency of our estimators. Simultaneously, we are lessening the degrees of freedom and, thus, the variety of statistical tests we can potentially employ. Our real gain from this is a drastic shrinking of the size of the data base we must manage. A vacuous gain also exists.

By using the per capita magnitudes for the values of our variables, we have not changed the underlying sample of individual observations, but have reduced the variability of the sample we are using for estimation. We have stripped the outlying individual observations of influence. The result is that the per capita magnitudes will be less dispersed around any expression we estimate, allowing us to appear to explain a larger proportion of the variation in the sample; that is, the magnitude of the coefficient of determination (R^2) is enhanced. This enhancement, however, is misleading since it is entirely due to our prior exercise of collapsing all the variations of individuals' observations in a city to a single scaler measure. Similarly, nonvacuously, and therefore much more importantly, by reducing the variation in the sample, we are reducing the standard errors of each estimated (and still unbiased) explanatory variable coefficient. As a consequence, we may be overstating the level of significance to be attached to these coefficients.

Yet another nonvacuous and altogether serious way exists for the estimates obtained from per capita data to be seriously misleading. The measurement unit one is using for any particular variable may differ from city to city. Thus, for example, one might be measuring cigarette consumption per capita in the equivalent of packs in one city and pounds in another. Consider the following simple algebraic argument.

Assume that a disaggregated dose-response expression for respiratory disease is to be estimated. Let this expression be given by:

$$C_{ij} = a_i + b_i p_{ij} + \varepsilon_{ij}, \quad \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, r \\ n \geq r \end{matrix} \tag{3.1}$$

where i refers to a particular pollutant, j to a particular individual, a and b are coefficients to be estimated, and ε is an error term having the customary properties. Per capita responses and doses are clearly:

$$C_j = \frac{\Sigma_i \, c_{ij}}{n} \tag{3.2}$$

$$P_j = \frac{\Sigma_i \, p_{ij}}{n} \tag{3.3}$$

With aggregation, the intercept and error terms are:

$$a = \frac{\Sigma_i \, a_i}{n} \tag{3.4}$$

$$E = \frac{\Sigma_i \, \varepsilon_{ij}}{n} \tag{3.5}$$

The aggregate relation is therefore:

$$Cj = a + bP_j + E_j, \tag{3.6}$$

where b, the coefficient of P_j, is apparently

$$b = \frac{\Sigma_i \, b_i}{n} \tag{3.7}$$

In other words, the per capita response depends on the exposures suffered by the n individuals. This perhaps seems reasonable, since (3.6) continues to be linear and includes an error term the expected value of which is zero for all c_{ij}.

Disregarding a and E_j, note, however, that both b and P_j are aggregated. Thus:

$$bP_j = \left(\frac{\Sigma_i \, b_i}{n}\right) \left(\frac{Pi}{n}\right) = \frac{\Sigma_i \, b_i \, p_{ij}}{n} \, , \tag{3.8}$$

and therefore

$$b = \frac{\Sigma_i b_i p_{ij}}{nP_j} \, . \tag{3.9}$$

Nothing goes awry if the dose-response functions are identical among sufferers. However, if they do differ, it is apparent from (3.8) that the value of the pollution exposure parameter, b, will be a <u>weighted</u> mean of the same parameter for the individual suffers. In particular, those sufferers having high responses will have a disproportionately strong influence upon a group's (e.g., a city) contribution to the value of the exposure parameter in (3.8). Similarly, those groups having low responses will have a disproportionately. weak influence. The conclusion is the rather dismaying one that the measure of responses, employing some group or aggregation of individuals as the fundamental unit of observation, can differ from one group to another. There could conceivably be as many unique measures employed as there are groups.

The preceding remarks refer to the prior aggregation of individual observations and the subsequent use of the aggregates for estimation purposes. Suppose we employ individual observations for estimation purposes, establish responses for the representative individual among these observations, and then use the presumedly representative responses of this representative individual to obtain an aggregate measure of total response; that is, we aggregate after rather than before estimation. The study of the morbidity effects of air pollution that follows readily lends itself to this treatment. Because it does so, we feel it worthwhile to caution the reader about the dangers this form of aggregation poses. We state the discussion in terms of demand functions although dose-response functions would serve equally well. Only because it is perhaps the most widely cited study to aggregate individual observations of air pollution damages, we employ Waddell (1974) as a basis for discussion.

Waddell (1974) first reviewed a collection of studies that had estimated marginal purchase price functions with respect to sulfur oxides and/or suspended particulates for eight different cities. Interpreting the values of the air quality parameters in these several studies as measures for the average household in each study of <u>equilibrium</u> marginal willingness to pay at given air quality states and with given demand functions for air quality, he selected a value within the range of these estimated values. By selecting this value within the range of values, he assumed that what was interpreted

as the equilibrium marginal willingness to pay was the same for all household in all cities.

Then, using the further assumption that this assumed equilibrium marginal willingness to pay was in fact the actual marginal willingness to pay for all air quality states, he multiplied the constant marginal willingness to pay by the number of households and the number of air quality states to obtain an estimate of aggregate national air pollution damages. That is, if b is the marginal willingness to pay and Q is an air quality state, Waddell (1974) calculated aggregate national air pollution damages, D, as

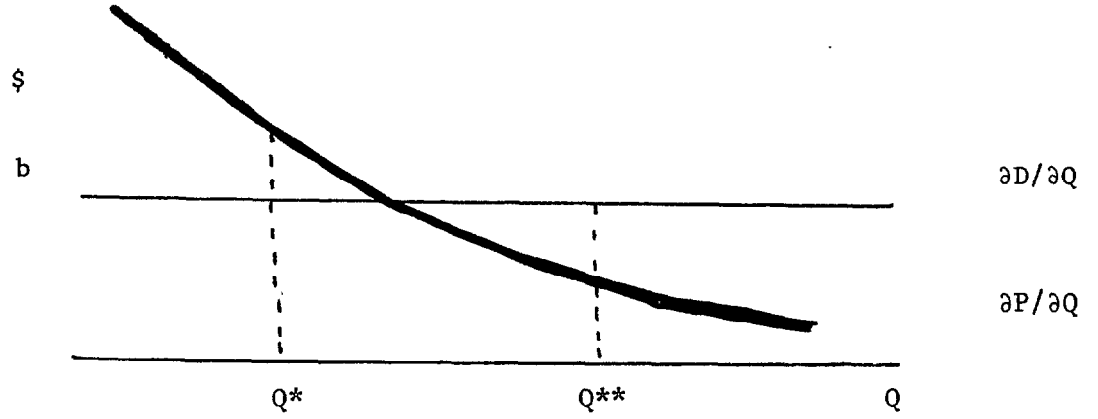
$$D = b\sum_{i=1}^{n}\Delta Q_i \qquad (3.10)$$

where the i's index households.

In effect, Waddell (1974) assumed that the decision problem of each and every household in each urban area of the country could be represented as depicted in Figure 3.1. In Figure 3.1, $\partial P/\partial Q$ is the marginal purchase price function and $\partial D/\partial Q$ is the function representing marginal willingness to pay for improvements in air quality. Since $\frac{\partial D}{\partial Q} = b$ is invariant with respect to changes in air quality, calculation of that willingness to pay for the household of Figure 3.1 involves only the multiplication of b by whatever change in air quality is postulated. Thus, the value to the depicted household of an improvement in air quality ($Q^{**} - Q^{*}$) is simply $b(Q^{**} - Q^{*})$. Given then that b is the same and invariant for all households, the sole distinction one need make among households in order to calculate aggregate national damages is to account for the location of each household on the Q axis.

Figure 3.1

Marginal Purchase Price and
Marginal Willingness-to-Pay



Among the more significant i.e., stronger assumptions in the aforementioned calculation procedure are the following. First, it is assumed in the procedure that the b's are invariant across households. By dropping this assumption, the immediately preceding expression becomes:

$$D = b_i \Sigma \Delta Q_i. \tag{3.11}$$

This would mean that differences in willingness to pay for improvements in air quality due to differences among households in such personal attributes as income, age, and degree of risk aversion to health effects would now be taken into account. Aggregation would then not entirely destroy knowledge about relative sufferer valuations of alternatives.

A further weakening of assumptions would occur if the marginal willingness to pay function is permitted to be nonconstant and even nonlinear. In this case, the above expression for D would be:

$$D = \Sigma_i \int_{Q^*}^{Q^{**}} b_i(Q_i)\, dQ_i \tag{3.12}$$

Clearly, this would be a very complex expression with which to calculate aggregate national air pollution damages. Not only are the marginal valuations of given air quality states permitted to vary among households but the responses of different households to similar variations in air quality are also permitted to differ. The sensitivity of the aggregation procedure to differences in the economic and air pollution circumstances of households would be greatly enhanced. Freeman (1974, pp. 81-82) lists several frameworks for constructing algorithms that might approximate this last expression for D.

The above discussion has been devoted to a single aggregation over individual households. It has been implicitly presumed that only a single class of air pollutants is relevant. Typically, however, estimates of national air pollution control benefits involve aggregation over multiple classes of pollutants as well as over households. On occasion, aggregation may, in addition, take place over time. Scaler estimates of the national benefits of air pollution control may thus involve two or three distinct types of aggregation, each of which embodies unique assumptions about the similarities among the units undergoing aggregation.

An additional decision problem, over and above the problem involving the manner in which the units in each type of aggregation are to be treated as similar, is thus introduced: one must choose thich type of aggregation is to be performed first in arriving at a scaler representing air pollution control benefits for households, for pollutants, and for time intervals. Moreover, in deciding how to perform the first aggregation, one must take into account how the aggregation for the second and third steps will be carried out. The order in which the aggregation is performed can make a difference in the estimate one obtains of aggregate national benefits.

Chapter IV

THE SIXTY-CITY EXPERIMENT

4.1 Objectives of the Experiment

Identification of substances in the environment which effect human health and accurate quantifications of their effects, is extremely difficult. Often there are multiple substances involved, there may be long latency periods before effects are seen, and the amount and time of exposureis often unknown. There are three general approaches to identifying such substances and quantifying their impact -- all more-or-less imperfect. In the first, laboratory animals are exposed to large doses of the suspect substance and, if effects appear, an effort is made to extrapolate them to the human population. The correct manner in which to execute the second step is not well extablished. The second approach is to develop aggregate cross-sectional data, usually for cities or standard metropolitan areas, on a number of variables which might be associated with mortality rates or illness rates and then to use regression analysis in order to discover statistically significant associations. A third approach is to gather very detailed data on individuals and to again use statistical analysis to determine the effect of various factors including environmental exposures on individualized measures of health status.

The purpose of the research reported in this chapter is to explore both the possibilities and limitations of the second approach mentioned above --aggregate epidemiology -- in the estimation of human dose-response functions which include exposure to air pollution. The principal advantage of the use of aggregated data on cities or metropolitan areas is quite simply the widespread availability and low cost of such data as opposed to data generated from animal experiments or collected on individual human beings through specialized surveys. However, the use of aggregated data creates a number of special problems.

First, one ideally wishes to estimate a dose-response relationship or function as shown in Figure 4.1. Based on a priori considerations one would suppose that for human populations, risk of death for an individual would be a function of medical care, age of the individual, the genetic endowment of the individual, the behavior of the individual--does he or she exercise, smoke, etc.--the diet of the individual, and exposures to possibly harmful substances or circumstances. However, aggregate epidemiology provides no data on individual risks or characteristics but only data for population characteristics as a whole. Thus, aggregate mortality rates in, for example a city are used as a proxy for risk of death in the

estimation of an individual dose-response function where it is implicitly assumed that individuals can be represented by the average individual in each city. Thus, in using the data set developed below for sixty U.S. cities to estimate a dose-response function of the form shown in Figure 4.1, it is implicitly assumed that each city represents one average individual. However, the list of assumptions required to allow such aggregation (all relationships must be perfectly linear, etc.) are not likely to be met in practice. Thus, one must recognize that estimated results are biased to an unknown extent by the very use of aggregated data.

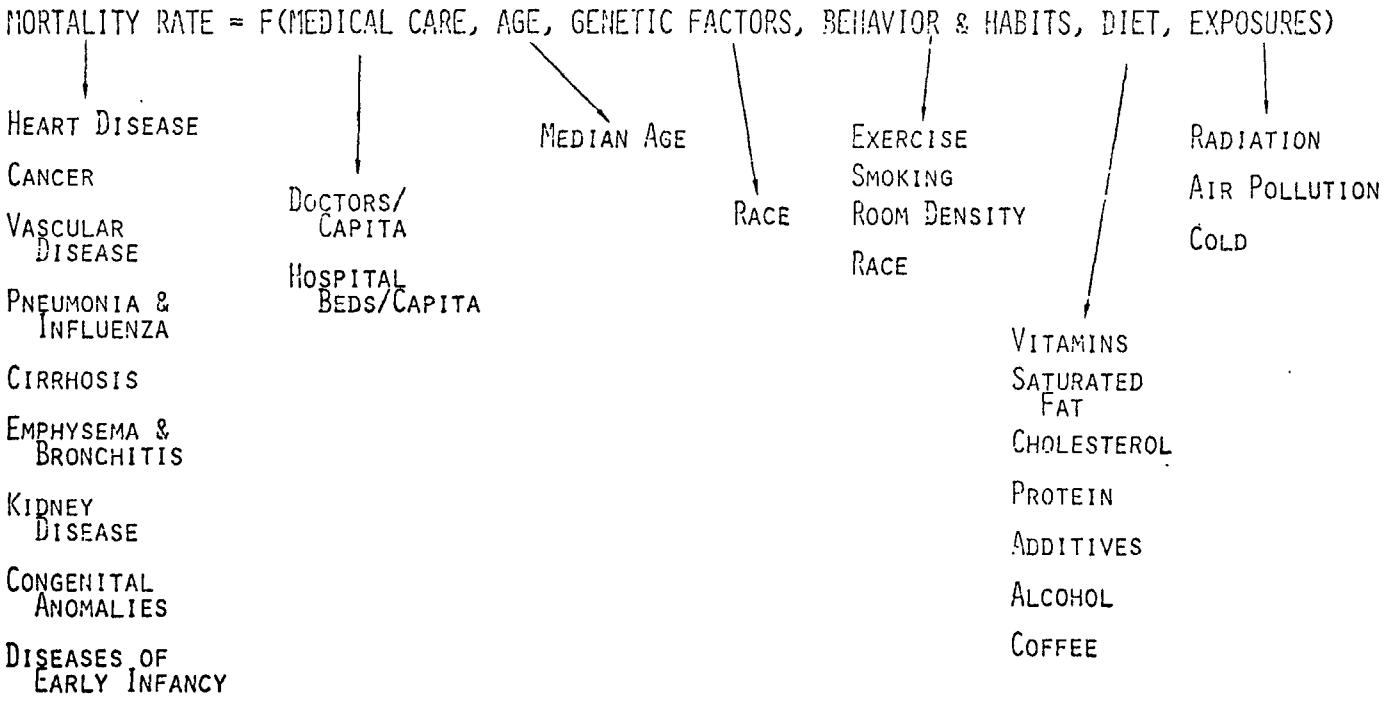
A second problem arises from the fact that aggregate epidemiology must rely on secondary data. Since the investigator must depend on data already collected, he cannot add a question to a survey nor can he vary the design of an animal experiment to test the importance of a new variable. In the past this problem has led to the exclusion of data on important variables such as smoking, diet and exercise from aggregate epidemiological studies [see, for example, Lave and Seskin (1977) and Schwing, et. al. (1974)] We have been able to gather some data -- not necessarily good data -- on both smoking and diet and as we show below, these are important omissions from previous studies. The current study still excludes any measure of exercise.

Finally, aggregate epidemiological studies are likely to suffer from a number of simultaneous equation biases. One of the most obvious concerns the effect of medical care. The existing epidemiological literature has failed to show any significant effect of medical care on human mortality rates. This counterintuitive result is easily explained. For example, in our sixty city sample, no effect of per capita doctors in each city can be shown on aggregate mortality rates for each city when simple regression techniques such as ordinary least squares are used. The explanation is that, although doctors most likely do reduce mortality rates (as shown below), people in cities with higher mortality rates have in turn more illness per capita and seek out more medical care, increasing the observed number of doctors in such communities. In other words, higher mortality rates create a greater demand for doctors. Thus, we have two offsetting effects--doctors reduce mortality, while mortality increases the demand for doctors--and simple regression analyses cannot untangle them. Several statistical techniques are available for coping with simultaneous equation problems.We use a very simple approach, two-stage least squares, a technique described in a little detail below.

A second simultaneous equation problem may arise because of multiple causes of death. Cities with high coronary death rates may likely have lower cancer death rates because people die of heart attacks before they have a chance to die of cancer. In this situation, factors which, for example, show up positively correlated with coronary disease may show up with a spurious negative correlation with cancer rates. This simultaneous equation problem is likely to work "through" the age variable in that median age is determined in part by mortality rates of individual diseases, while, in turn, age is used to explain mortality rates. We therefore have also employed two-stage least squares on the age variable, but with no impact on the estimated equations so these results are not reported here.

Figure 4.1

Hypothetical Human Dose-Response Function



An alternative approach to the problem which we do not employ, is use of age-specific mortality rates.

A third possible source of simultaneous equation bias occurs because people make voluntary choices over location. Migration in and out of our sixty city sample is effectively disregarded. People may, for example, contract an air pollution related disease and, on a physician's advice, move from a highly polluted area to an unpolluted area, only then to die. A false negative association between air pollution and pollution-related mortality might then be shown. Although in the past we have included a net migration variable [see Kneese and Schulze (1977)] which was statistically significant, we have excluded such a variable in this analysis because it defies interpretation in a dose-response function context.

Table 4.1 summarizes the objectives and limitations of the current study and to some extent those of aggregate epidemiology in general. We now turn to development of methodology for estimating the value of reducing health risks and for the effect of medical care on human health. This latter section focuses on the role of exonomics in aggregate epidemiology.

4.2 Value of Life Vs. Value of Safety

The direct costing approach employed by economists for evaluating the mortality costs of diseases which result from environmental exposures is straightforward but difficult to quantify fully [see, for example, Kneese and Schulze, 1977]. First, the population at risk must be known. Second, the increased risk of mortality associated with environmental exposures must be quantified either through epidemiology or through extrapolation from animal experiments. Third, the amount of money or the value that individuals place on safety (avoiding risk of death) must be known. Multiplying these three values together then gives an approximation of the incremental benefits of reducing such exposures. This cost or benefit is not in any way related to a "value of life" which is most likely unmeasureable, but rather focuses on a concept of the value of safety (alternatively "cost of risk") to individuals where risks are statistically small.

Table 4.1

Objectives and limitations

Purposes of Study Are:

(1) To explore methodology for isolating an aggregate human dose-response function.

(2) to add medical inputs.

(3) to add diet.

(4) to add smoking.

(5) to account for simultaneous equation bias where possible including:

(a) demand for doctors.

(b) multiple possible causes of death.

The Study Fails to Account for:

(1) simultaneous equation problems caused by migration.

(2) exercise.

(3) radiation.

(4) Biases introduced by estimating an aggregate as opposed to individual dose-response function.

Economists in the past have attempted to value human life as the sum of the present value of future earnings over an individual's lifetime [see Lave and Seskin, 1970 and 1977]. This approach, however, is no longer viewed as acceptable. In the first place, it assumes that the value of life can, in fact, be measured -- a point certainly open to debate. Second, it implies that the lives of children, housewives, retired and other unemployed individuals are worth less than the lives of employed heads of households.

Two measures can be used to value safety or risk to life which are based on the economic concepts of equivalent variation (EV) and compensating variation (CV). An EV measure of the value of life is the amount of money an individual would pay to escape from or prevent certain death; in theory, a rational individual would part with all his available wealth to save his life. CV, in contrast, measures the compensation required to induce an individual to accept voluntarily a situation where the probability of death is increased. As the probability of death approaches unity, the CV measure can be taken as an estimate of the value the individual places on his life. Logically, though, the value of life measured this way must be infinite, because as the probability of death approaches certainty, the probability of enjoying any compensation offered (and thus the value of the compensation) approaches zero. Thus, neither EV (which requires coercion) nor CV (which makes the value of life immeasurable) provides a wholly satisfactory way of extimating the dollar costs of mortality in realy world situations that involve risk. An eleboration of the CV concept, however, can provide a useful measure of the compensation necessary to induce an individual to accept a slight increase in the probability of death.

Mishan (1971) was the first to distinguish between the concept of cost of risk, which is ethically appealing, and earlier efforts to value human life based on lost earnings, which as a methodology, has strange and intuitively objectionable features. The latter measure of the "value" of human life has now been rejected by economists both on theoretical and, to some extent, on ethical grounds. Thaler and Rosen (1975), using wage differentials between jobs varying in the level of job-associated risk of death, were the first to estimate explicitly the value of safety. In other words, workers in high risk jobs receive higher wages and a value of safety can be inputed by examining risk-related wage differentials. Unfortunately, however, their study dealt with a high risk class of individuals. The Thaler and Rosen (1975) estimate suggests that in current dollars a small reduction in risk over a large number of individuals which saves one life is worth about $340,000. Another study [Blumquist (1977)], which examines seat belt use, suggests that the figure might be $260,000. This study first estimates how people value their own time and then imputes a value of safety from the amount of time a sample of individuals spent in buckling up seat belts. These results may be biased downward because individuals seem to perceive risks differently when an element of personal control such as driving an automobile is involved. Finally, Smith (1975) in a study similar to Thaler and Rosen (1975) has suggested that, for a more typical population and for job-related risks, the figure may exceed $1,000,000. Clearly, the cost of risk is not precisely known, and perhaps will never be, since attitudes -- risk preferences -- presumably can change over time, between groups, and can even vary in different situations. But, we at

least have a range of values with which to make order-of magnitude estimates of the costs of environmental risks. This range of values does not, however, overlap the value-of-life estimates based on lost earnings. For example, Lave and Seskin (1977) use a value in the thirty to forty thousand dollar range for a life lost. The Thaler and Rosen (1977) value of safety is, for example, about an order of magnitude larger than the Lave and Seskin (1977) lost earnings number.

The theoretical basis of a value of safety or cost of risk concept can be shown briefly as follows: Assume that an individual has a utility function, U(W), where utility is an increasing function of wealth, W. If risk or death is Π, expected utility is $(1-\Pi)U(W)$. If we hold expected utility constant, we have $(1-\Pi)U(W)$ = constant, and the total differential of this equation is:

$$-U(W)d\Pi + (1-\Pi)\, U'(W)dW = 0 \qquad (4.1)$$

where the prime denotes differentiation. Holding utility constant then implies that the increase in wealth (or income) necessary to offset an increase in risk is:

$$dW/d\Pi = U/[U' (1-\Pi)]. \qquad (4.2)$$

This is the compensating variation measure of the cose to an individual attributable to an increased risk of death. Analysis of the last expression can be simplified if we assume a constant elasticity of utility with respect to wealth, η, such that U(W) = η and consequently $\eta = \frac{dU}{dW}\frac{W}{U}$. Then (4.2) can be rewritten as:

$$dW/d\Pi = W/[\eta\ (1-\Pi)]. \qquad (4.3)$$

The right hand side of (4.3) suggests several interesting points about the value of safety or cost of risk. First, if we assume that the elasticity of utility is less than one, people are risk averse. This in turn implies that since the risk of death is positive $(\Pi>0)$ that $(dW/d\Pi)>W$. In other words, if an individual is risk averse, his life, in terms of the risk premium necessary to get him to accept risk, is worth more to him than his wealth. Second, from (4.3), as wealth increases, the risk premium required to accept an increase in risk voluntarily, $dW/d\Pi$ must increase with age, ceteris paribus. Thus, one would expect older people to act in a more risk averse manner than younger individuals (require greater compensation to voluntarily take a risky action), both because of increased income and because of increased initial age-related risk of death.

This model contrasts for a number of reasons with the value of lost earnings approach previously used in economic analysis. First, if lost income itself is the measure, the "value of life" measured through lost earnings obviously cannot exceed wealth [see Conley, 1977]. Second, increased wealth will increase the lost earnings measure as well as the cost of risk measure. However, the cost of risk measure may not increase proportionately if a different utility function is used. Third, the lost earnings measure must decrease with age at some point as individuals get

older because the expected remaining earnings must decrease, while the cost of risk, as we argued above, will likely increase. Finally, it is clear from (4.3) that as Π approaches unity, $dW/d\Pi$ approaches infinity. In other words, the compensation required to induce an individual to accept a certainty of death voluntarily is infinite. The lost income measure has no similiar property. Nevertheless, the implication is that small increases in risk may be valued in terms of compensation required to induce individuals to accept such risks voluntarily. Individuals, of course, rationally accept small risks on a daily basis; presumably on the basis of some monetary or psychic return.

Given the analysis above, the current methodology of multiplying value of safety numbers times experimentally or epidemiologically determined enviromental risks can then be justified as follows: assuming a utility function U(W) where W is wealth, if risk of death is Π, the marginal cost of risk, as derived earlier, is $(dW/d\Pi)\bar{U} = U/U'(1-\Pi)$, where $\bar{U}$ is a constant utility level. If risk, Π, is a function of pollution, X, where utility functions are identical for N individuals, one would wish to maximize expected utility,

$$N[1-\Pi(X)]U(W), \tag{4.4}$$

subject to a constraint on total wealth, $\bar{W}$, or income of society

$$\bar{W} - NW - C(X^{o}-X) = 0 \tag{4.5}$$

which is allocated to individual wealth, assumed identical for purposes of exposition, (NW), and costs of controlling environmental pollution from the initial level X^{o}, $[C(X^{o}-X)]$. Noting that $\Pi_{X} > 0$, and $C' > 0$, the first order conditions are (where λ is the multiplier on (4.5) and L denotes the Lagrangian):

$$\partial L/\partial W = N(1-\Pi)U' - N\lambda = 0$$

$$\partial L/\partial X = -N\Pi_{X}U - \lambda C' = 0$$

These imply:

$$N\cdot [U/U'(1-\Pi)]\Pi_{X} = C' \tag{4.6}$$

or that the number of individuals, N, times the marginal cost of risk, $[U/U'(1-\Pi)]$, times the marginal effect of pollution on risk, Π_{X}, equals the marginal cost of control, C'. Clearly, this model abstracts from many welfare threoretic problems but it does imply that estimation of the left hand side of (4.6) as suggested at the beginning of this section is a legitimate approximation of the incremental benefits of environmental control.

In summary, the direct costing of mortality has the advantage of focusing attention on one positive output of environmental agencies which has clear economic value -- safety. It is important, however, to distinguish between the value of safety to consumers which does have measurable economic value -- environmental agencies may be viewed as selling safety

to the public -- as opposed to techniques which claim to measure the value of human life. Benefit-cost arguments for environmental programs should and can rest on demonstrable increases in public safety delivered at costs comparable to what the public is willing to pay for safety, not on claims as to the value of human life. However, the assessment of the risk of mortality associated with environmental exposures such as air pollution--whether based on animal experiments or epidemiological studies -- remains difficult and uncertain and is central to the direct costing methodology. Surprisingly, perhaps, the authors feel there is likely to be less professional debate as to the economic measure of the dollar value of safety than as to the quantification of environmental health effects. We now turn to the possible role of economic analysis in the epidemiology of air pollution.

4.3 A Methodological Basis: Does Economics Matter?

The question posed above could be rephrased "does rational human behavior matter in the estimation of dose-response functions?" Economists would certainly answer in the affirmative; individuals are likely to respond to illness with numerous ameliorative measures. Clearly, such measures must be accounted its a properly specified dose-response function is to be estimated. What follows is a simplified economic model of human behavior in response to health risks which in turn allows specification of appropriate statistical techniques for estimating a human dose-response relationship.

Let Π denote risk of death for an individual where that risk can be reduced by medical care which we denote D, synonymous with our empirical measure, doctors per capita. Thus, risk can be written as a function, $\Pi(D)$, where $d\Pi/dD = \Pi' < 0$. If the price of medical care is p and income is denoted Y, then utility, U, can be written $U(Y - pD)$, a function of income net of expenditures on medical care, pD. In an uncertain world, economic analysis assumes that an individual will choose to maximize expected utility -- the odds of remaining alive $(1-\Pi)$ times the utility level U -- or

$$[1 - \Pi(D)]\, U(Y - pD), \tag{4.7}$$

so the first order condition for the quantity of medical care chosen when rearranged is:

$$\frac{U}{(1-\Pi)U'} = \frac{P}{-\Pi'} \tag{4.8}$$

The term on the left-hand side of (4.8) is easily recongizable from section 4.2 above as the marginal value of safety (or compensation required to voluntarily accept a small increase in risk), while the term on the right is the marginal cost of increased safety through medical care. Thus, this model of human behavior implies that an individual will choose a level of medical care which equates his or her marginal value of safety to the marginal cost of reducing risk through medical care. Of course, an individual's perception of risk and of the ability of medical care to reduce risk of death may be imperfect. However, from (4.8) it is easy to show that individuals who are more risk averse, i.e., those with a large marginal

value of safety, will seek more medical care than those who are less risk averse.

An explicit set of functional forms will simplify interpretation. Let us again (as in Section 4.2) assume a constant elasticity of utility with respect to income, η, so $U = (Y - pD)^{\eta}$. Also assume a linear (approximate) dose-response relationship, $\Pi = \Pi_o + \Pi' D$, where $\Pi' < 0$ is now a fixed coefficient and Π_o is a positive constant. Equation (4.8) can then be written as:

$$D = \left(\frac{\eta}{\Pi'}\right) - \left(\frac{1}{\Pi'}\right)\Pi + \left(\frac{1}{P}\right)Y \tag{4.9}$$

which is a demand equation for medical care. If we take the supply price of medical care to be fixed $P = P^*$ (infinitely elastic supply of medical care), the individual demand for medical care, doctors per capita for example, is then a linear increasing function of total risk Π, since $\left(\frac{-1}{\Pi'}\right) > 0$, and of income Y, since $\frac{1}{P^*} > 0$. Of course, we wish, as our principal objective for policy purposes, to estimate the dose-response function:

$$\Pi = \Pi_o + \Pi' D; \tag{4.10}$$

in particular, we wish to obtain an unbiased estimate of Π', the effect of medical care on mortality and of the effect of other variables such as air pollution. However, any attempt to directly estimate (4.10) is doomed to failure. This occurs because the equation specified for statistical estimation (equivalent to 4.10 where α_o and α_1 are parameters for estimation)

$$\Pi = \alpha_o + \alpha_1 D + \mu_\Pi \tag{4.11}$$

has a disturbance term μ_Π which is not independent of D. In other words, μ_Π is correlated with D. This is easy to show if we specify the demand equation for doctors (equivalent to 4.9 above with parameters β_o, β_1 and β_2) as stochastic:

$$D = \beta_o + \beta_1 \Pi + \beta_2 Y + \mu_D \tag{4.12}$$

as well, with a disturbance term μ_D. Now suppose some factor (random) embodied in μ_Π causes Π to rise in (4.11). But if Π rises, by (4.12), D must increase since, from (4.9), $\beta_1 > 0$ and D, through β_1, depends on Π. Thus, D depends on μ_Π through (4.12) and D and μ_Π are correlated. Now, if in estimating (4.11) this correlation is not accounted for, not only will estimates of α_1 and α_0 be biased, but if we had included other factors which affect morality such as diet or pollution in (4.11), coefficients on these variables would be biased as well. It is also true that if simultaneous equation biased is

present and not accounted for, it becomes possible that the estimated effect of medical care, α_1, will appear not significantly different from zero or even of the wrong sign (note we assume that $\alpha_1 < 0$; that doctors reduce mortality).

We can break the dependence of D on μ_Π by first substituting (4.10) into (4.9), or (4.11) into (4.12), to obtain a reduced form equation for medical care,

$$D + \frac{\eta}{2\Pi'} - \frac{\Pi^o}{2\Pi'} + \frac{1}{2P^*}Y \text{ or } D = \gamma_o + \gamma_1 Y + \mu_r \tag{4.13}$$

where μ_r is the disturbance term in the reduced form. This equation can be legitimately estimated since the income variable is exogenous, determined outside the relevant system of equations, and the endogenous variables D and II, those determined within the system, do not appear on the right-hand side of (4.13). Now, if we estimate (4.13) and obtain unbiased estimates of the two coefficients $\gamma_o = (\frac{\eta}{2\Pi'} - \frac{\Pi}{2\Pi'})$ and $\gamma_1 = (\frac{1}{2P^*})$ we can use these along with data on income, Y, to generate a new variable, estimated medical care, D, where

$$D = \gamma_o + \gamma_1 Y. \tag{4.14}$$

Note that this new variable, D, generated from data on Y does not depend on μ_Π and can be used instead of actual data on D to estimate a dose-response function:

$$\Pi = \alpha_o + \alpha_1 D + \mu_\Pi \tag{4.15}$$

This estimated equation gives a consistent estimate of α_1 or Π'. In fact, if the hypothesis that doctors are both important and effective in reducing mortality rates is correct, α_1 should show up negative and significantly different from zero as estimated in (4.15). Note, however, that if individuals perceive that doctors are effective, they will have a strong incentive to seek medical help when ill, thus making a direct least squares estimate of the effect of medical care as specified in (4.11) impossible.

The procedure we have outlined above, two-stage least squares, has been used successfully in many instances to resolve simultaneous equation problems and has the advantage of requiring minimal additional data. In general, if an unbiased estimate of a structural equation (one equation is a simultaneous system) is desired, one need only use ordinary least squares to estimate each of the endogenous variables as a function of all of the exogenous variables in the model (estimate a set of reduced form equations). Then, using the data on the exogenous variables, an estimated data set for each of the endogenous variables is created. Consistent structural equations can then be obtained by replacing each endogenous variable

on the right hand side of a structural equation by its estimated equivalent using ordinary least squares.

4.4 The Sixty-City Data Set: Selection of Variables

In this section, we describe the data set itself and also examine some properties of the data with special emphasis on collinearity and consequent implications on the variety and kinds of hypotheses which can be appropriately tested.

Tables 4.2-4.5 present a listing of the variables available for analysis along with the year of the variable, units, mean, standard deviation (S.D), and sources for the data by number, where the number refers to the listing of sources in Table 4.6. Table 4.2 includes total mortality rate calculated from 1970 data on mortality by city divided by 1970 census population. Disaggregated mortality data by disease category -- heart, vascular, pneumonia and influenze, emphysema and bronchitus, cirrhosis, nephritis and nephosis, congenital anomalies, early infant diseases, and cancer -- were also collected for 1970, and divided by 1970 census population to develop mortality rates; exceptions are the congential anomolies and early infant disease categories which were divided by the number of births in each city for 1970 in order to define an appropriate mortality rate. Mortality data for 1970 were chosen because reliable city population estimates are available for that year as opposed to more recent data requiring use of non-census year city population estimates in place of actual data. The disaggregation of total mortality by disease may not be appropriate. However, only data on city mortality was available, as indicated in Table 4.2.

Table 4.3 describes per capita dietary data by city for the years 1955 and 1965. The procedure used to construct the dietary data sets was somewhat involved. Food consumption estimates were first constructed for each of the 60 cities, using data on a sample of about 3,000 urban households, distributed among eight income brackets, for four regions of the United States, collected by the Department of Agriculture for 1955 and 1965. The results are regionally-specific weighted averages of consumption of various foods by families in each income bracket, multiplied by the fraction of each city's population in each income bracket. Data for specific dietary factors were then generated by multiplying the consumption rates of 49 foods by their respective concentrations of a given substance. A number of additional variables are available from the Department of Agriculture for 1965 as opposed to 1955. These include total protein, total fats, and total carbohydrates. As such, these variables provide a better indication of broad dietary patterns as opposed to the 1955 data set.

Table 4.4 describes our data on socioeconomic, geographic, and smoking variables. The socioeconomic and geographic variables were chosen for their consistency for estimating the aggregate dose-response function hypothesized in previous sections. Both the income and education variables are hypothesized to enter the demand equation for medical care, <u>not</u> the dose-response function. We must therefore employ the two-&age least squares estimation technique outlined above. Doctors per capita was chosen as the best available indication of available medical care, both in terms of

Table 4.2

Mortality Variables

Variable		Year	Units	Mean	S.D.	Sources
Mortality Variables						
M070	Total Mortality	1970	deaths/1000 pop.	11.283	2.161	(18) (6)
HA70	Heart Disease	1970	"	4.216	1.078	(18) (6)
VA70	Vascular Disease	1970	"	1.566	0.395	(18) (6)
PN70	Pneumonia & Influenza	1970	"	0.375	0.114	(18) (6)
EM70	Emphysema & Bronchitis	1970	"	0.178	0.059	(18) (6)
CI70	Cirrhosis	1970	"	0.238	0.106	(18) (6)
NE70	Nephritis & Nephrosis	1970	"	0.058	0.027	(18) (6)
C/B%	Congenital Anom/Births	1970	%	0.473	0.105	(18) (6)
I/B%	Early Infancy/Births	1970	%	1.294	0.333	(18) (6)
CA70	Cancer Mortality	1970	deaths/1000 pop.	1.958	0.402	(18) (6)

Table 4.3

Dietary Variables

Variable		Year	Units	Mean	S.D.	Source
Dietary Variables						
NTRI	Nitrites in Food	1955	g/yr/cap	1.27	0.14	(2) (4) (27)
NTRA	Nitrates in Food	1955	g/yr/cap	69.86	9.05	(2) (4) (27)
SFAT	Saturated Fatty Acids*	1955	g/yr/cap	16220.00	874.65	(2) (3) (4)
PROT	Protein*	1955	g/yr/cap	26557.00	1314.00	(2) (3) (4)
CHOL	Cholesterol*	1955	g/yr/cap	234.81	6.98	(2) (3) (4)
CVIT	Vitamin C**	1955	g/yr/cap	16.96	1.46	(2) (3) (4)
CALO	Calories	1955	kcal/yr/cap	1149.7	56.27	(2) (3) (4)
COFF	Coffee	1955	kg/yr/cap	5.83	.70	(2) (3) (4)
ALCO	Alcohol (S value)	1955	$/yr/cap	17.30	6.06	(2) (3) (4)
XPRO	Total Protein	1965	g/yr/cap	39845.	706.46	(28)
XFAT	Total Fats	1965	g/yr/cap	57512.	1795.7	(28)
XCAR	Carbohydrates	1965	g/yr/cap	123490.	3623.0	(28)
XASA	Ascorbic Acid	1965	mg/yr/cap	42281.	2364.2	(28)
6NTI	Nitrites in Food	1965	g/yr/cap	1.14	.16	(4) (28) (27)
6NTA	Nitrates in Food	1965	g/yr/cap	52.87	2.47	(4) (28) (27)
6SFT	Saturated Fatty Acids*	1965	g/yr/cap	16315.	976.3	(3) (4) (28)
6PRO	Protein*	1965	g/yr/cap	28128.	1603.4	(3) (4) (28)
6CHL	Cholesterol*	1965	g/yr/cap	219.9	5.80	(3) (4) (28)
6CAL	Calories	1965	kcal/yr/cap	1171.1	27.63	(3) (4) (28)
6CVT	Vitamin C**	1965	g/yr/cap	18.65	1.3	(3) (4) (28)
6COF	Coffee	1965	kg/yr/cap	5.40	.18	(3) (4) (28)
6ALC	Alcohol ($ value)	1965	$/yr/cap	25.97	6.45	

* Includes only animal products.

** Includes only vitamin C content for fruits and vegetables eaten fresh.

Table 4.4

Social, Economic, Geographic, and Smoking Variables

Variable		Year	Units	Mean	S.D.	Sources
Social, Economic, Geographic						
MDOC	Medical Doctors	1970	M.D.'s/100,000	162.8	54.2	(19)
IN69	Median Income	1969	$/yr/Household	10763.	1060.	(6)
EDUC	Education	1969	%>25 yrs w/H.S. diploma	55.3	7.4	(6)
DENS	Crowding In Homes	1969	%>1.5 persons/ room	0.022	0.013	(8)
COLD	Cold Temperatures	1972	#days temp < 0 0°C.	86.9	47.7	(9)
NONW	Nonwhite Population	1969	Fraction	0.226	0.154	(6)
MAGE	Median Age of Population	1969	Years	28.82	2.74	(6)
Smoking Variables						
C156	Cigarettes	1956	packs/yr/cap†	183.52	26.66	(22) (4)
C168	Cigarettes	1968	packs/yr/cap††	165.80	23.25	(7) (1)

† Data for states, 1960 census population used.

†† Data for states, 1970 census population used.

Table 4.5

Air Pollution Variables

Variable		Year	Units	Mean	S.D.	Sources
Air Pollution						
SU66	Sulfate	1966	$\mu g/m^3$	10.67	5.44	(20)
AM66	Ammonium	1966	$\mu g/m^3$	1.15	1.42	(20)
NI66	Nitrates	1966	$\mu g/m^3$	1.96	0.68	(20)
PA66	Suspended Particulates	1966	$\mu g/m^3$	114.83	33.97	(20)
NO69	Nitrogen Dioxide	1969	ppm	0.076	0.034	(11)
PA70	Suspended Particulates	1970	$\mu g/m^3$	102.30	30.11	(13)
SO70	Sulfur Dioxide	1970	$\mu g/m^3$	26.92	22.20	(13)
NI70	Nitrate, annual mean	1971-73	$\mu g/m^3$	3.13	0.92	(14)
NI90	Nitrate, 90th %-tile con.	1971-73	$\mu g/m^3$	5.21	1.80	(14)
SU70	Sulfate, annual mean	1971-73	$\mu g/m^3$	10.65	4.01	(14)
SU90	Sulfate, 90th %-tile con.	1971-73	$\mu g/m^3$	17.69	7.62	(14)
LEAD	Lead	1970	$\mu g/m^3$	1.33	0.54	(16)
CO74	Carbon Monoxide	1974	mg/m^3	11.86	3.50	(15)
BETA	Beta Radioactivity	1966	PC_i/m^3	0.261	0.091	(20)

preventative and ameliorative care. Alternative variables such as hospital beds per capita were judged inferior, in that underutilization of hospital facilities is a common problem and adjustments for utilization factors would prove troublesome. Also, if one assumes a less than perfectly elastic supply of medical care, the doctors per capita variable is an appropriate supply side variable in that it reflects patient loads for doctors in a particular city. The possible importance of age and cold temperatures in a dose-response relationship are clear. However, the nonwhite and crowding variables may be more difficult to interpret. The nonwhite variable would ideally control to some extent for genetic variations in the population. Obviously, however, this variable may well proxy for education, poverty, habits, etc. Similarly, crowding would ideally be an indication of possible contagion but may really proxy for poverty, old age, or even race. Thus, the role of these variables should be interpreted with great care.

Cigarette consumption was estimated from cigarette tax revenues for each state in which a sample city was located; the result is thus a statewide average that includes rural populations. Per capita cigarette consumption was estimated using the total state population over 16 years of age both for 1956 and 1968. It should be noted that both our dietary and our smoking variables are quite crude, reflecting problems in utilizing secondary data. However, the possible importance of their effects on human health may justify use of even these measures. We also attempted to develop a measure of total exposure to radiation, collecting data on beta radioactivity in air, terrestrial gamma radiation, and cosmic ray exposures, but have to this point been able to account for only about half of the average individual annual dose associated with medical exposures. As a result, no suitable total exposure variable is available at this time.

Table 4.5 presents the air pollution variables available for testing. All data are annual geometric means for each city unless otherwise specified. Unfortunately, hydrocarbon data was only available for about ten of our sample cities and are excluded for this reason. Finally, as noted above, Table 4.6 presents our data sources.

In summary, data available for testing include: (1) 1970 mortality rates for total mortality and for major disease categories; (2) data on dietary patterns for 1955 and 1965; (3) data on medical doctors and socioeconomic factors for 1970 or a nearby year; (4) data on smoking patterns for 1956 and 1968; and (5) data on air quality for each city for 1970 or a nearby year.

Since only 60 observations are available, we must obviously select a subsample of the available explanatory variables for hypothesis testing. To allow straightforward statistical tests of the significance of estimated coefficients, it is necessary to make the selection of included variables *a priori* rather than testing each of the variables in various combinations for significance and excluding some on the basis of relative significances. Techniques do exist for testing significance where pre-testing has been employed but the standard t-statistic is no longer applicable.

The first problem in specifying the final data set is a decision on including lagged variables. Given a highly mobile U.S. population, the

Table 4.6

Sources of Data

(1) Advisory Commission on Intergovernmental Relations, State and Local Significant Features and Suggested Legislation, 1972, Table 120, 1970.

(2) U.S. Department of Agriculture, Household Food Consumption Survey, 1955, Reports No. 2-5.

(3) _____, Composition of Foods: Raw, Processed and Prepared, Watt, Bernice K., and Merrill, Annabell L., Agricultural Handbook No. 8, 1968.

(4) U.S. Department of Commerce, Bureau of the Census, U.S. Census of the Population: 1960.

(5) _____, Cross Migration by County: 1965-1970, Current Population Reports Series P-25, No. 701, issued May 1977.

(6) _____, U.S. Census of Population: 1970, Vol. 1-50.

(7) _____, State Tax Collections: 1968, Series GF 68 No. 1, Tables 7, 9.

(8) _____, 1970 Census of Housing by State.

(9) _____, National Oceanic and Atmospheric Administration, Environmental Data Service, Climatological Data, National Summary: Annual 1972, Vol. 23, No. 13, Asheville, North Carolina.

(10) U.S. Environmental Protection Agency, Natural Radiation Exposure in the United States: 1972, Report No. ORP/SID 72-1, Table A-1, 1974 Reprint.

(11) _____, Air Quality Criteria for Nitrogen Dioxide, No. AP-84, Tables 6-10, January 1971.

(12) _____, Chemical Analysis of Interstate Carrier Water Supply System, PB-257600/7BE April, 1975.

(13) _____, Air Quality Data - 1970 Annual Statistics, EPA-450/2-76-019, October 1976.

(14) _____, Air Quality for Nonmetallic Inorganic Ions, 1971 through 1974: From the National Air Surveillance Networks, EPA-600/4-77-003, January, 1977.

(15) _____, Air Quality Data - 1974 Annual Statistics, EPA 450/2-76-011, August, 1976.

(16) U.S. Environmental Protection Agency, Air Quality Data for Metals 1970 through 1974: From the National Air Surveillance Networks, EPA-600/4-76-041.

Table 4.6

(continued)

(17) Department of Health, Education and Welfare, Public Health Service, National Center for Health Statistics, Vital Statistics of the United States: 1960.

(18) _____, Vital Statistics of the United States: 1970.

(19) _____, Health Manpower -- A County and Metropolitan Area Data Book.

(20) _____, National Air Pollution Control Administration, Air Quality Data from the National Surveillance Network and Contributing State and Local Networks, 1966 Edition.

(21) _____, Vital Statistics of the United States: 1972.

(22) Tobacco Tax Council, Cigarette Taxes in the United States, 1956, Table 15.

(23) Directory of Medical Specialists, 1960-71, Marquis - Who's Who, Inc., Chicago, Illinois.

(24) Adams, John A., et. al., eds., The Natural Radiation Environment II, Proceedings of the Second International Symposium on the Natural Radiation Environment, Houston, Texas, August 7-11, 1972.

(25) Hickey, John, et. al., The Development of an Engineering Control Research and Development Plan for Carcinogenic Materials, U.S. Government Printing Office, Washington, D.C. (1977 in press).

(26) Pazand, Reza, Environmental Carcinogenesis - An Economic Analysis of Risk, PhD. Dissertation, The University of New Mexico, July 1976.

(27) White, Jonathan W., Jr., "Relative Significance of Dietary Sources of Nitrate and Nitrite," Journal of Agricultural and Food Chemistry, Vol. 23, No. 5 (1975), Table VI, p. 890.

(28) U.S. Department of Agriculture, Household Food Consumption Survey, 1965-66, Reports No. 7-10 and Reports No. 13-16.

Table 4.7

Simple Correlation Matrix for 1965 Diet Variables

	XPRO	XFAT	XCAR	XASA	6NTI	6NTA	6SFT	6PRO	6CHL	6CVT	6COF	6ALC
XPRO	1.00	0.46	-0.01	0.53	-0.64	0.43	0.16	0.29	0.67	0.88	-0.14	0.70
XFAT	0.46	1.00	0.85	-0.17	0.34	0.74	-0.62	-0.41	0.50	0.12	-0.09	-0.28
XCAR	-0.01	0.85	1.00	-0.33	0.66	0.58	-0.66	-0.46	0.33	-0.31	-0.22	-0.71
XASA	0.53	-0.17	-0.33	1.00	-0.69	-0.16	0.86	0.93	0.58	0.79	-0.59	0.58
6NTI	-0.64	0.34	0.66	-0.69	1.00	0.20	-0.66	-0.67	-0.36	-0.77	0.25	-0.93
6NTA	0.43	0.74	0.58	-0.16	0.20	1.00	-0.47	-0.31	0.46	0.20	0.16	-0.13
6SFT	0.16	-0.62	-0.66	0.86	-0.66	-0.47	1.00	0.96	0.26	0.55	-0.39	0.54
6PRO	0.29	-0.41	-0.46	0.93	-0.67	-0.31	0.96	1.00	0.51	0.63	-0.53	0.50
6CHL	0.67	0.50	0.33	0.58	-0.36	0.46	0.26	0.51	1.00	0.64	-0.48	0.21
6CVT	0.88	0.12	-0.31	0.79	-0.77	0.20	0.55	0.63	0.64	1.00	-0.16	0.81
6COF	-0.14	-0.09	-0.22	-0.59	0.25	0.16	-0.39	-0.53	-0.48	-0.16	1.00	0.03
6ALC	0.70	-0.28	-0.71	0.58	-0.93	-0.13	0.54	0.50	0.21	0.81	0.03	1.00

question becomes, "do people now living in a city represent the same sample as individuals who lived in cities 14-15 years before (the lags on smoking and diet, respectively)?" If the answer is no, and if people carry their dietary and smoking characteristics with them as they move, the most recent available data is likely to better represent long-term dietary and smoking patterns of individuals in a particular city. For these reasons, in this study, we use the available data closest to 1970 throughout. However, it may well be, for diseases with long lags such as cancer, that lagged variables are superior in any case. The real answer is, of course, to account properly for mobility -- a near impossibility when using aggregate data.

The second consideration in specifying variables for inclusion is multicollinearity, Typicall, multicollinearity problems can be avoided if the simple correlations between explanatory variables are less than 0.4 to 0.6. Tables 4.7 and 4.8 present simple correlation matrices for 1965 diet and air quality variables, respectively.

Table 4.7 shows that a very high level of collinearity is probably present among dietary variables. It appears so high, in fact, that the problem becomes one of finding a set of dietary variables which is sufficiently non-collinear to allow reasonable estimation of individual coefficients. Perhaps the broadest indicators of dietary patterns are the total protein (XPRO), total fat (XFAT) and total carbohydrate (XCAR) variables. Protein and fat will tend to indicate high consumption of meat. and nuts, while high carbohydrate consumption will indicate consumption of grains, fruits, vegetables, and refined sugars. However, total carbohydrate and total fat have a correlation coefficient of 0.85, too high to allow inclusion of both variables in an estimated equation. However, if we replace total fats (XFAT) with animal fat (6SFT), a good proxy for consumption of saturated fats, the correlation between fat (now animal fat) and carbohydrates drops to 0.66, which although still high, will likely cause less difficulty. Thus, we include these three diet variables as broad indicators of dietary patterns where, however, it must be clearly recognized that the estimated coefficients on these variables may well include the effect partially or totally of a number of other highly collinear dietary variables. For example, total protein (XPRO) has a correlation with cirrhosis, one might justifiably doubt that a causal relationship exists between protein and cirrhosis as opposed to one between slcohol and cirrhosis.

Table 4.8 suggests that multicollinearity may well be a problem within the air quality data set as well. Given previous research (see, for example, Lave and Seskin, 1977), the air quality measures of most concern are those for NO_2, SO_2, sulfate, and particulates, so we focus on these variables here. However, our measures of SO_2 and sulfate for 1970, the year of the mortality data, are highly collinear -- S070 and SU70 have a simple correlation of 0.74 -- so any separation of their relative importance is likely impossible. As a result, we use SO_2 (S070) as a proxy variable for both pollutants. Note also that among the included air pollution variables, NO_2 (N060), particulates (PA70), and SO_2 (S070), collinearity problems may exist with respect to ammonium, carbon monoxide and lead (some correlations greater than or equal to 0.4). Since we exclude these variables here (as